Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **471,** 2174–2186 (2017) doi:10.1093/mnras/stx1772
Advance Access publication 2017 July 14

The VLT LBG redshift survey – VI. Mapping H I in the proximity of $z \sim 3$ LBGs with X-Shooter

R. M. Bielby,[1⋆] T. Shanks,[1] N. H. M. Crighton,[2] C. G. Bornancini,[3,4] L. Infante,[5] D. G. Lambas,[3,4] D. Minniti,[6,7,8] S. L. Morris[1] and P. Tummuangpak[9]

[1]*Durham University, South Road, Durham, DH1 3LE, United Kingdom*
[2]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[3]*Instituto de Astronomía Teórica y Experimental (IATE, CONICET-UNC), Laprida 854, X5000BGR, Córdoba, Argentina*
[4]*Observatorio Astronómico (OAC), Universidad Nacional de Córdoba, Laprida 854, X5000BGR, Córdoba, Argentina*
[5]*Instituto de Astrofísica, Facultad de Física, Centro de Astroingenneiría, Pontificia Universidad Católica de Chile Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile*
[6]*Departamento de Ciencias Fisicas, Universidad Andres Bello, Republica 220, Santiago, Chile*
[7]*Vatican Observatory, I-00120 Vatican City State, Italy*
[8]*The Millennium Institute of Astrophysics (MAS), Santiago, Chile*
[9]*Department of Physics, Khon Kaen University, Khon Kaen, 40002, Thailand*

Accepted 2017 July 12. Received 2017 July 12; in original form 2016 October 5

ABSTRACT

We present an analysis of the spatial distribution and dynamics of neutral hydrogen gas around galaxies using new X-Shooter observations of $z \sim 2.5$–4 quasars. Adding the X-Shooter data to our existing data set of high-resolution quasar spectroscopy, we use a total sample of 29 quasars alongside ∼1700 Lyman Break Galaxies (LBGs) in the redshift range $2 \lesssim z \lesssim 3.5$. We measure the Ly$\alpha$ forest auto-correlation function, finding a clustering length of $s_0 = 0.081 \pm 0.006\,h^{-1}$ Mpc, and the cross-correlation function with LBGs, finding a cross-clustering length of $s_0 = 0.27 \pm 0.14\,h^{-1}$ Mpc and power-law slope $\gamma = 1.1 \pm 0.2$. Our results highlight the weakly clustered nature of neutral hydrogren systems in the Lyα forest. Building on this, we make a first analysis of the dependence of the clustering on absorber strength, finding a clear preference for stronger Lyα forest absorption features to be more strongly clustered around the galaxy population, suggesting that they trace on average higher mass haloes. Using the projected and 2-D cross-correlation functions, we constrain the dynamics of Lyα forest clouds around $z \sim 3$ galaxies. We find a significant detection of large-scale infall of neutral hydrogen, with a constraint on the Lyα forest infall parameter of $\beta_{\rm F} = 1.02 \pm 0.22$.

Key words: intergalactic medium – galaxies: kinematics and dynamics – large-scale structure of Universe – cosmology: observations.

1 INTRODUCTION

The relationship between gas and galaxies is a crucial component of galaxy formation models. Star formation cannot be sustained without the supply of gas available to galaxies from their surroundings and understanding the flow of gas in and out of galaxies is imperative for a complete understanding of galaxy formation. Gas in the interstellar medium (ISM) provides the reservoir from which the star-formation process is fuelled, but once this process is under way, winds from stars and supernovae begin to drive the interstellar gas outward (Heckman, Armus & Miley 1990; Lehnert & Heckman 1996; Wilman et al. 2005). The supernovae in particular may produce high-velocity winds that are able to act on large scales, driving gas and metals out of the galaxy and into the intergalactic medium, where ≈80 per cent of baryons reside at $z \sim 3$ (Bi & Davidsen 1997). Outflows are thought to heat and enrich gas within the galaxy halo (i.e. within the galaxy virial radius), also referred to as the circum-galactic medium (CGM), reducing the amount of gas in the galaxy available for star formation. A similar feedback process is expected to occur as a result of active galactic nuclei activity, driven by powerful jets. Simulations have indicated that these outflows may play a crucial role in galaxy evolution (e.g. Springel & Hernquist 2003; Springel, Di Matteo & Hernquist 2005; Bower et al. 2006; Oppenheimer & Davé 2006; Sijacki et al. 2007; Oppenheimer & Davé 2008; Oppenheimer

⋆ E-mail: richard.bielby@durham.ac.uk

et al. 2010; Schaye et al. 2010; Shen, Wadsley & Stinson 2010; Davé, Oppenheimer & Finlator 2011; Hopkins, Quataert & Murray 2011, 2012; Hopkins et al. 2014).

Tracing the distribution and the dynamics of the gas presents a significant challenge to observational astronomy however. Although some progress has been made in recent years in attempting to trace the CGM gas via faint emission (e.g. Steidel et al. 2011; Wisotzki et al. 2016), the prime method for probing both the CGM and the inter-galactic medium (IGM) remains that of identifying absorption features in Quasi-Stellar Object (QSO) spectra. Given the ubiquity of hydrogen in the Universe, Lyman series Hydrogen absorption features offer insights into a range of environments, ranging across galaxy voids (e.g. Tejos et al. 2012), filamentary structures (e.g. Tejos et al. 2016) and warm gas structures potentially tracing galaxy outflows or the intra-group medium (e.g. Morris & van den Bergh 1994; Hoffman et al. 1999; Chen & Prochaska 2000; Pachat et al. 2016; Bielby et al. 2017).

With the information from QSO sightline absorbers in hand, it is important to explore how these absorbers relate spatially to the galaxy population and large-scale structure in general. Early statistical analyses of the distribution of absorbers with respect to the galaxy population showed tentative evidence for some clustering or 'clumpiness' of absorbers around galaxies at $z \lesssim 0.3$ (Bahcall et al. 1992; Morris et al. 1993). This early work was developed further, extending to larger samples and higher redshifts, by subsequent studies (e.g. Lanzetta et al. 1995; Chen et al. 1998, 2001). Morris & Jannuzi (2006) detected a significant correlation between H I absorbers and galaxies at separations of $\lesssim$1.5 Mpc, albeit weaker than the galaxy–galaxy auto-correlation. They found their results to be consistent with the absorbing gas and the galaxies coexisting in dark matter filaments and knots, consistent with predictions from models of galaxy formation.

At higher redshift, Adelberger et al. (2003) presented the first analysis of the cross-correlation between H I in quasar sightlines and Lyman-break galaxies (LBGs) at $z \approx 3$. Their results showed a clear increase in absorption in quasar sightlines within $\approx 5\,h^{-1}$ Mpc of the positions of LBGs in their survey. Further to this, at very small separations ($\lesssim 500\,h^{-1}$ kpc), they found a spike in the transmission profile suggestive of a lack of H I gas in the immediate proximity of the $z \sim 3$ galaxies. The authors suggested that this might be the result of galaxy winds driving the gas away. However, the result at these scales was based on only three galaxies in their sample of ∼800 and it has proven difficult to re-create such a feature in simulations of star-forming galaxies. Following this work, the same group presented a similar analysis based on galaxies at $z \sim 2$ (Adelberger et al. 2005). The same approach was taken using a larger sample of galaxies and no such spike in the absorption profile close to galaxies was observed, although some uncertainty remained given the differing redshift ranges of the two studies.

At $z \lesssim 1$, further studies developed the low redshift absorber cross-correlation measurements (e.g. Ryan-Weber 2006; Wilman et al. 2007; Shone et al. 2010; Tejos et al. 2014). Ryan-Weber (2006) and Tejos et al. (2014) are of particular interest in that they probe the 2-D cross-correlation function, investigating evidence for dynamical effects on the absorber distribution around galaxies. Indeed two papers find somewhat conflicting results: Ryan-Weber (2006) show prominent extensions in the correlation function along the line of sight (to ≈400–600 km s^{-1}), claiming these to be the result of the intrinsic galaxy–absorber velocity dispersion, whilst Tejos et al. (2014) place an upper limit of $\lesssim$100 km s^{-1} on the intrinsic velocity dispersion.

Following the same methods as Adelberger et al. (2003, 2005), Crighton et al. (2011, hereafter Paper II) presented the first analysis of the H I-galaxy cross-correlation using a wide area survey of LBGs at $z \sim 3$ using the VIMOS instrument on the Very Large Telescope (VLT): namely the VLT LBG Redshift Survey (VLRS). This analysis was based on a sample of ∼1000 $z \sim 3$ galaxies surrounding seven quasar sightlines at $z \sim 3$. They found a deficit in Lyα transmission within ∼5 Mpc (comoving) in agreement with Adelberger et al. (2003) and Adelberger et al. (2005), but no evidence for the upturn in average transmission seen by Adelberger et al. (2003).

At a redshift range of $2 < z < 3$, Rudie et al. (2012) and Rakic et al. (2012) present analyses of the distribution of H I gas around galaxies using ∼800 galaxies from the Keck Baryonic Structure Survey (KBSS). Rudie et al. (2012) use Voigt profile fitted H I absorption lines in 15 QSO spectra to analyse the gas distribution, finding evidence for infalling gas at large scales and peculiar velocities of ≈260 km s^{-1} at scales of $\lesssim$400 kpc. Similarly, Rakic et al. (2012) use pixel-optical-depth analysis using the same data and reach equivalent conclusions. Although benefiting from high densities of galaxies, the KBSS fields are constrained to relatively small scales of $\approx 6h^{-1}$ Mpc (comoving) limiting the efficacy with which they may discern the Kaiser effect due to infalling material (Kaiser 1987).

Following this, Tummuangpak et al. (2014, hereafter Paper IV) presented the latest galaxy-H I cross-correlation results of the VLRS survey, using an updated sample of 17 QSOs and ≈2000 $z \sim 3$ LBGs. This was built on the work of Paper II again showing the flux decrement around LBGs and a lack of any transmission spike at small scales. Paper IV presented a first full model-based analysis of redshift-space distortions (RSDs) of the gas around galaxies, finding a low-velocity dispersion consistent with the redshift uncertainties on the galaxy positions and a tentative measurement of gas infall at large scales. Their observational results were presented alongside the analysis of a simulated volume from the Galaxies-Intergalactic Medium Interaction Calculation (GIMIC) hydrodynamical simulation that matched their observational data.

In this paper, we present an analysis of the relationship between gas and galaxies at redshifts of $z \approx 3$ based on spectroscopic observations of LBGs and a combination of new moderate resolution quasar spectra from the VLT X-Shooter and high-resolution quasar spectra. The X-Shooter data offer improved statistical power over previous papers, affording improved constraints on the clustering of the Lyα forest around LBGs, in particular studying the dependence of clustering on absorber strength (i.e. connecting halo mass to absorber strength). Significantly, the expanded data set also provides improvements in measuring the large-scale infall of neutral hydrogen gas towards the star-forming galaxy population. In Section 2, we present the observations, detailing the data reduction for the X-Shooter data and giving an overview of the QSO and galaxy data used. Section 3 presents our analysis of the Lyα forest auto-correlation function incorporating the X-Shooter data, whilst in Section 4 we present the galaxy-Lyα forest cross-correlation analysis. We discuss our results in terms of the absorber dynamics and the relationship between absorbers and the underlying dark matter distribution in Section 5. In Section 6, we present our conclusions.

This is the sixth in a series of papers presenting the VLRS. Bielby et al. (2011, hereafter Paper I) presented the initial sample of ∼1000 $z \approx 3$ galaxies combined with an analysis of galaxy clustering. Paper II analysed the gas–galaxy cross-correlation based on this first sample of $z \approx 3$ galaxies. Bielby et al. (2013, hereafter Paper III) presented an updated sample totalling ∼2000 LBGs and presented an analysis of galaxy clustering at $z \approx 3$. Paper IV built

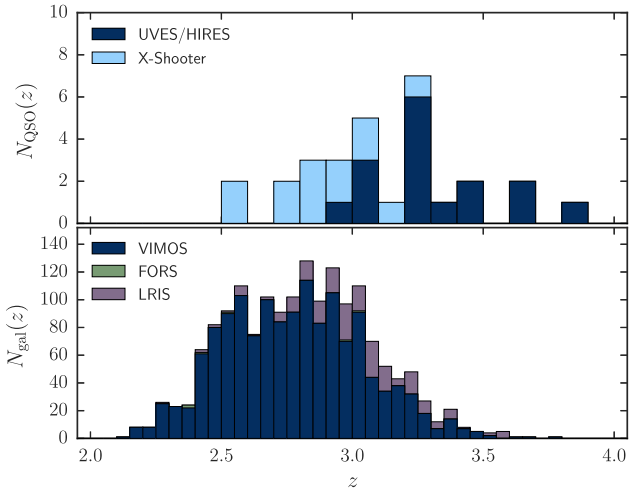

Figure 1. Top panel: The stacked redshift distributions of the QSO sample observed with the high-resolution instruments HIRES and UVES (dark blue-filled histogram) and the QSO sample observed with the moderate-resolution instrument X-Shooter (light blue-filled histogram). Lower panel: The stacked redshift distribution of the 1665 spectroscopically observed galaxies included in this analysis (i.e. with redshifts lower than the redshift of any QSO in our sample of 29 that lies within 30 arcsec), incorporating galaxies observed using VIMOS (dark blue histogram), FORS (green histogram) and LRIS (purple histogram).

on the early VLRS results using GIMIC to analyse the results in the context of detailed simulations of galaxies and IGM. Finally, Bielby et al. (2016, Paper V) present a survey of Lyman-α emitters in a subset of the VLRS fields, although these data are not included in the analysis presented here (as we keep a focus on the relatively consistent selection of LBGs).

Throughout this paper, we use a cosmology given by $H_0 = 100\,h\,{\rm km\ s^{-1}}$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $\sigma_8 = 0.8$. Distances are quoted in comoving coordinates in units of h^{-1} Mpc, unless otherwise stated.

2 OBSERVATIONS

2.1 Survey overview

The data used here cover 14 separate fields, each centred on a bright $z \gtrsim 3$ QSO. Below we provide details of the QSO sightline and galaxy redshift data available in those fields.

2.1.1 *High-resolution QSO spectra*

In the case of all but one of the fields, archive high-resolution spectra are available from either VLT UVES or Keck HiRES for the central bright quasar (the exception being Q2359-0653). In addition, three further quasars ([WHO91] 0043-265, LBQS 0302-0019, and Q212904.90-160249.0) within the primary fields have archival high-resolution spectra available. In total, this provides 16 QSOs with high-resolution spectroscopy of the Lyα forest. The redshift distribution of these is given by the dark blue histogram in Fig. 1.

We have performed the reduction of the available high-resolution Echelle spectra for these quasars and this is described in full in Paper II and Paper IV. An overview of these quasars is given in Table 1.

For the purposes of our analysis in this paper, we wish to avoid the inclusion of Lyman Limit Systems (LLS) and Damped Lyα absorber (DLA) systems (due to our pixel-based analysis), as well as regions in the spectra that have not been observed. Taking individual features in the quasar spectra, the spectrum of Q0301-0035 is masked in the wavelength range 4460 Å $< \lambda <$ 4520 Å, which contains a gap in the data. Also, the HIRES spectrum of J1201+0116 contains two gaps (4420 Å $< \lambda <$ 4540 Å and 4790 Å $< \lambda <$ 4850 Å), which are masked. The UVES spectrum of Q2348-011 contains a gap in the wavelength range 4510 Å $< \lambda <$ 4630 Å, as well as DLAs at ∼4160 Å and ∼4400 Å. LLSs and DLAs are also present and masked in the spectra of WH091 0043-265 ($\lambda \sim$ 4640 Å), J0124+0044 ($\lambda \sim$ 4950 Å) and Q212904.90-160249.0 ($\lambda \sim$ 4650 Å).

2.1.2 *VLT X-Shooter spectra*

Paper II presented a survey of $R < 22$ quasars within the VLRS fields, with low-resolution spectra of 295 quasars observed using AAOmega at the AAT. Here, we present VLT X-Shooter moderate resolution spectra of a selection of these, which overlap with the VLRS galaxy sample. In total, we use 15 QSOs observed using the X-Shooter instrument as part of ESO programs 085.A-0327 and 087.A-0906 (2 of which also have high-resolution data available). The list of quasars used here is given in Table 1, whilst the redshift distribution is shown by the pale blue component of the histogram in Fig. 1.

The observations were performed in NOD mode with individual exposure times of 694, 695 and 246 s with the UVB, VIS and NIR arms, respectively. For quasars with magnitudes of $R \leq 20$, two exposures were made in the UVB arm, two with the VIS arm and six with the NIR arm. Quasars fainter than $R = 20$ were observed with double the number of exposures used for the brighter quasars. Slit widths of 1.0, 1.2 and 1.2 arcsec were used for the UVB, VIS and NIR arms, respectively, giving resolutions of $R = 4350$, $R = 6700$ and $R = 3890$ in each arm. Standard flux observations were made using the spectrophotometric standard stars GD71, LTT7987 and EG 131.

The X-Shooter spectra were reduced using the ESO X-Shooter pipeline package version number 1.4.6 and the ESOREX command line reduction tool. We followed the standard reduction procedure as outlined in the X-Shooter Pipeline User Manual. All of the X-Shooter spectra were flux calibrated using the observed spectrophotometric stars.

2.1.3 *Quasar continuum fitting*

In order to quantify the absorption along the line of sight to the observed quasars, we first need an estimate of the intrinsic quasar continuum and broad-line flux, $f_c(\lambda)$. This is performed using a suite of bespoke quasar absorption line python tools (QALPY[1]) as developed and used by Crighton et al. (2010); Shone et al. (2010) and Paper II for the analysis of quasar spectra.

For each quasar, we use the script FITCONTINUUM to estimate the intrinsic continuum, which follows the methods of Young et al. (1979) and Carswell et al. (1982). Each QSO spectrum is first divided into wavelength intervals, in which the mean and standard deviation of the observed flux are calculated. A sigma-clipping process is then employed to iteratively reject the most aberrant pixels, until the remaining pixels show an approximately Gaussian distribution; the mean flux of these pixels is then taken to be the

[1] Available online at https://github.com/nhmc/barak

Table 1. The full list of QSOs present in this study, observed with VLT UVES, VLT X-Shooter and Keck HIRES. In total, we use 29 QSOs: 14 observed solely at high resolution, 13 observed solely at moderate resolution, and a further 2 observed at both high and moderate resolutions. The resolution and S/N per pixel in the Lyα forest (LyαF) range are given for each QSO (note pixel scales for UVES/HIRES and X-Shooter are $\approx$0.04 and $\approx$0.15 Å, respectively). Sources for the galaxy data are given as V – VIMOS, L – LRIS and F – FORS.

Name	RA	Dec (J2000)	R (Vega)	z	Instrument	Resolution (km s^{-1})	LyαF S/N	$N_{\rm gal}$	Galaxy data
Q000033.06+070716.1	00:00:33.06	+07:07:16.1	19.6	2.86	X-Shooter	44	7.5	110	V
Q000127.48+071911.8	00:01:27.48	+07:19:11.8	20.7	2.87	X-Shooter	44	3.4	110	V
Q000137.67+071412.2	00:01:37.67	+07:14:12.2	20.8	2.99	X-Shooter	44	5.6	109	V
Q2359+0653	00:01:40.60	+07:09:54.0	18.4	3.23	X-Shooter	44	28	97	V
Q000234.97+071349.3	00:02:34.97	+07:13:49.3	20.6	2.60	X-Shooter	44	5.0	50	V
LBQS 0041-2638	00:43:42.79	−26:22:10.2	18.3	3.05	X-Shooter	44	15.4	165	V
Q0042-2627	00:44:33.95	−26:11:19.9	18.5	3.29	HIRES	6.7	4.0	159	V
[WHO91] 0043-265	00:45:30.47	−26:17:09.2	18.3	3.44	HIRES	6.7	13.6	139	V
J0124+0044	01:24:03.78	+00:44:32.7	17.9	3.81	UVES	7.5	20.9	72	V
Q0201+1120	02:03:46.7	+01:11:34.4	20.1	3.610	HIRES	6.7	6.9	12	L
Q0256-0000	02:59:05.6	+01:00:11.2	18.2	3.364	HIRES	6.7	28	37	L
Q030241.61-002713.6	03:02:41.61	−00:27:13.6	20.1	2.81	X-Shooter	44	5.0	151	V,L,F
SDSS J030335.42-002001.1	03:03:35.45	−00:20:01.1	19.9	2.72	X-Shooter	44	6.3	133	V,L,F
LBQS 0301-0035	03:03:41.05	−00:23:21.9	17.6	3.23	HIRES	6.7	25	113	V,L,F
LBQS 0302-0019	03:04:49.86	−00:08:13.5	17.5	3.29	HIRES	7.5	18	108	V,L,F
QSO B0933+289	09:33:37.2	+01:28:45.3	17.8	3.428	HIRES	6.7	10	37	L
HE0940-1050	09:42:53.50	−11:04:25.9	16.6	3.05	UVES	7.5	45	310	V
J1201-0116	12:01:44.37	+01:16:11.7	17.4	3.20	HIRES	6.7	19	63	V
Q1422+2309	14:24:38.1	+01:22:56.0	.	3.620	HIRES	6.7	59	69	L
Q212904.90-160249.0	21:29:04.90	−16:02:49.0	19.2	2.90	UVES	6.7	6.7	89	V
PKS2126-158	21:29:12.15	−15:38:40.9	17.3	3.27	UVES	7.5	72	126	V
Q2231-0015	22:34:09.00	+00:00:01.7	17.3	3.02	UVES/X-Sh	7.5/44	40	80	V,F
Q2233+136	22:36:27.2	+01:13:57.1	18.7	3.210	HIRES	6.7	8.0	29	L
Q234919.94-010727.0	23:49:19.94	−01:07:27.0	20.8	2.75	X-Shooter	44	2.5	110	V,F
SDSS J234921.56-005915.1	23:49:21.56	−00:59:15.2	19.9	3.09	X-Shooter	44	6.3	184	V,F
Q234958.23-004426.4	23:49:58.23	−00:44:26.4	21.0	2.58	X-Shooter	44	2.9	78	V,F
Q2348-011	23:50:57.88	−00:52:09.9	18.7	3.01	UVES/X-Sh	7.5/44	15	174	V,F
Q235119.47-011229.2	23:51:19.47	−01:12:29.2	20.1	2.94	X-Shooter	44	4.6	157	V,F
Q235201.36-011408.2	23:52:01.36	−01:14:08.2	20.4	3.12	X-Shooter	44	5.5	184	V,F

continuum level in that wavelength bin. An initial estimate of the continuum is then made by performing a cubic spline interpolation across the continuum levels estimated in each wavelength bin. Final adjustments are then made to the fit by hand where the continuum fit appears poor (i.e. around damped Lyα systems and emission lines). For further details of the script, we refer the reader to Paper II.

2.2 Galaxy data

The galaxy data used in this analysis consist exclusively of galaxies selected using the Lyman Break method in optical filters (U, B and R or u, g and r) and covering a redshift range of $2 \lesssim z \lesssim 3.5$. These cover a magnitude range of $23 \lesssim R \lesssim 25.5$. Whilst the full VLRS galaxy redshift data also incorporate Lyα emitters at $R \gtrsim 25.5$ (Bielby et al. 2016), we limit the sample in this paper to just the LBG selected population in order to use a consistently selected galaxy sample. In total, we use 1665 galaxies lying at redshifts coeval with the Lyα forest of a background quasar. These galaxies are split between three observational sources: VLT VIMOS, VLT FORS and Keck LRIS. Each of these samples is described in further detail below.

2.2.1 *VLT VIMOS data*

The current catalogue of LBGs from the VLRS has been presented in Paper III. It consists of a sample of 2147 spectroscopically confirmed galaxies, which were initially selected using the Lyman-break technique across nine of the QSO fields discussed above. The galaxy sample was observed with low-resolution VIMOS spectroscopy using the LR_Blue prism ($R \sim 180$) under ESO programs 075.A-0683 (PI: D. Minniti), 077.A-0612 (PI: T. Shanks), 079.A-0442, 081.A-0418 and 082.A-0494 (PI: L. Infante). Redshifts were measured based on fits to the Lyα and ISM spectral features, followed by corrections for the velocity offsets of these lines, giving velocity accuracies of $\sim$360 km s^{-1} (Paper III). Each galaxy identification was given a confidence flag from 0.5–1.0, where Paper III shows that $\sim$60 per cent of $Q = 0.5$ objects are correct based on multiple observations of single objects.

The numbers of VLRS $z \sim 3$ galaxies in each field are given in Table 1, whilst the redshift distribution is shown in Fig. 1. Taking the full sample of QSOs used in this study, we present the distribution of VLRS LBG – QSO sightline pairs as a function of separation, s (in units of h^{-1} Mpc), in Fig. 2 (blue histogram). The total number of available galaxies within the requisite redshift range (i.e. between

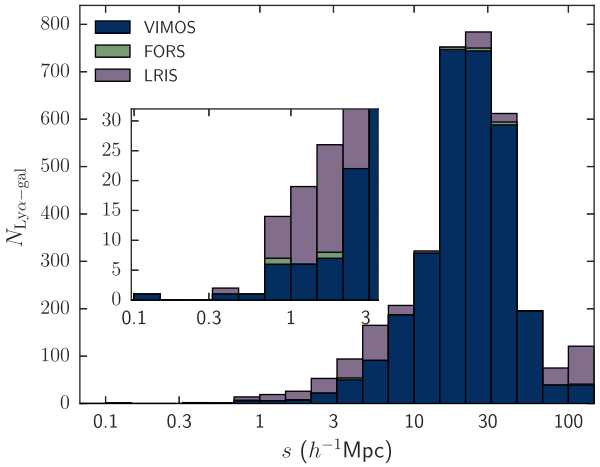

Figure 2. Number of LBG-sightline pairs ($N_{Ly\alpha\text{-LBG}}$) as a function of pair separation, s. The filled blue histogram shows the VIMOS detected galaxies, the green histogram the additional FORS detected galaxies, and purple the added LRIS galaxies (with the three histograms stacked to form the total). The inset shows a zoom in of the pair count histograms at small separation.

the Lyα and Lyβ emission redshifts of background QSOs) to be included in our analysis is 1437.

2.2.2 *VLT FORS data*

In addition to the VIMOS galaxy data, we also incorporate follow-up spectroscopy of galaxies close to quasar sightlines using the VLT FORS instrument. These consist of re-observations of some galaxies to get more accurate redshifts plus observations of previously unobserved LBG candidates. The FORS observations were made close to three of the quasars used in this study – Q0301-0035, Q2231-0015 and Q2348-011 – and were observed as part of ESO program 087.A-0627. The velocity accuracy for the FORS observations is estimated to be ≈150–180 km s^{-1}.

The FORS data in these fields comprise reliable redshifts for 11 $z \sim 3$ galaxies, 3 of which were originally identified using the VLRS VIMOS data and 8 of which lie within the requisite redshift range. These additional galaxies primarily lie $0.6 < s < 3h^{-1}$ Mpc from QSO sightlines in their respective fields, as shown in Fig. 2 (green histogram).

2.2.3 *Keck LRIS data*

Six of the fields are sampled by the Keck LRIS data of Steidel et al. (2003) and we incorporate these data into our sample. The data consist of 328 $z \sim 3$ galaxy redshifts around the six available background QSOs, adding 220 additional galaxies within the Lyα–Lyβ wavelength range of background QSOs to our total sample. The LRIS galaxy redshifts have been corrected for systematic shifts between features and have a velocity uncertainty of ≈250 km s^{-1}.

We note that one of the Steidel et al. (2003) fields overlaps with one of the VLRS fields around the background QSO, LBQS 0302-0019. We find three galaxies observed by both surveys and find consistent redshifts in each case. For the purposes of this work, we use the Keck LRIS redshift in each of these three cases.

The Keck LRIS data are shown in the sightline-LBG pair counts in Fig. 2 (purple histogram) and predominantly add to the counts in the range $0.4 < s < 10\,h^{-1}$ Mpc.

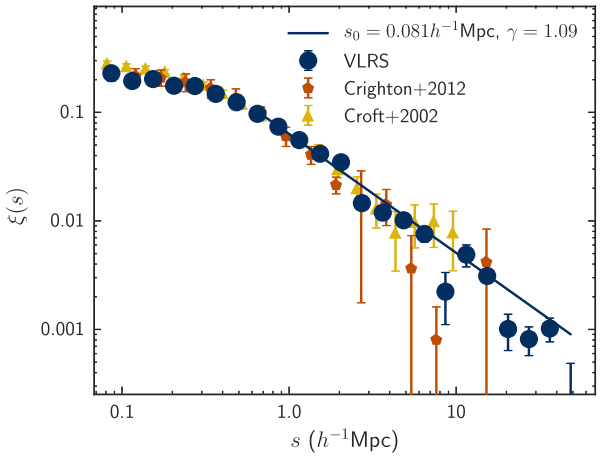

Figure 3. The Lyα forest auto-correlation result for the 24 quasar sightlines studied here (dark filled circles). Pale filled squares show the result when limited to only the π-direction. The solid line shows the best fit to our full result at separations of $s > 1.0\,h^{-1}$ Mpc. Our previous results (Paper II) are given by the pentagon points, whilst the literature results of Croft et al. (2002) are shown by filled triangles.

3 AUTO-CORRELATION OF THE LYα-FOREST

Using our 29 moderate and high-resolution QSO spectra, we first measure the Lyα-forest auto-correlation function. This constrains the mass distribution of the forest in preparation for then analysing the cross-correlation with the galaxy population.

We perform the auto-correlation following the pixel-to-pixel method used in Paper II. Using the normalized (transmitted) flux, $T = f/f_c$, we evaluate:

$$\delta_F = \frac{T}{\overline{T}(z)} - 1 \quad (1)$$

where $\overline{T}(z)$ is the mean normalized flux. If we discount high column density systems, the quantity δ_F traces the underlying mass density fluctuations, δ_m, in a relatively linear manner (e.g. McDonald, Miralda-Escudé & Cen 2002). The measured analytical form of $\overline{T}(z)$ at $z \approx 3$ is given by the following (McDonald et al. 2000):

$$\overline{T}(z) = 0.676 - 0.220(z - 3) \quad (2)$$

The auto-correlation function, $\xi_F(r)$, is then given by:

$$\xi_F(\Delta r) = \langle \delta_F(r)\delta_F(r + \Delta r) \rangle \quad (3)$$

In this way, we calculate ξ_F using the individual normalized flux pixel values in each quasar sightline. To prevent the high-resolution data dominating our results in terms of sheer number of pixels, we re-sample the data (using a mean-binning) to match the X-Shooter pixel scale of 0.15 Å/pixel.

The result of the auto-correlation calculation using the 29 quasar sightlines is shown in Fig. 3 (dark filled circles). This updated result is consistent with the previous VLRS results presented in Paper II (orange pentagons) and Paper IV, although we show the result to larger scales than in these previous papers. The use of moderate resolution data (and the necessary resampling of the high-resolution data to match) does not show any evidence of significantly affecting the results when compared to Paper II (which is based on a subsample of our own data). The yellow triangles show the results of Croft et al. (2002), where we take the average of their subsamples D and E (see their table A6) in order to best match the mean QSO redshift of our sample. Again, the results are consistent where the

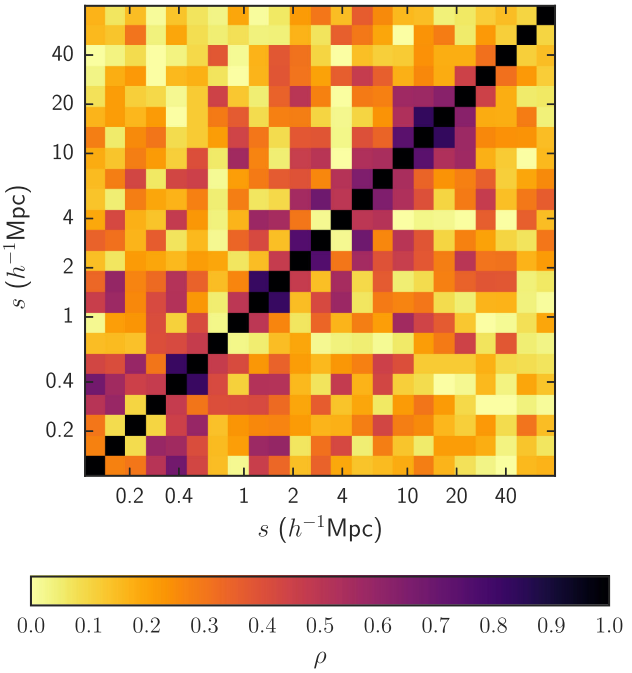

Figure 4. The covariance coefficient measured for the Lyα forest 1D auto-correlation is shown in Fig. 3.

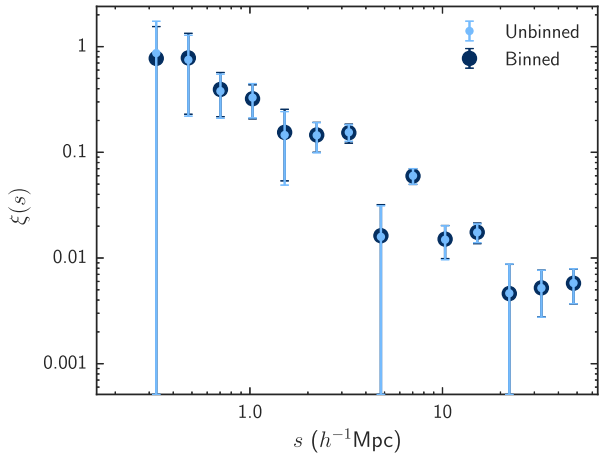

Figure 5. The result of the cross-correlation analysis using only the 11 background quasars from VLRS with high-resolution spectra. The large circular points show the result based on the native pixel scale of the high-resolution data (0.03 Å/pixel), whilst the small circles show the data re-sampled to the X-Shooter pixel scale (0.15 Å/pixel).

scales probed overlap (and $s \gtrsim 0.4\,h^{-1}$ Mpc), suggesting that the measurements are robust.

We evaluate the covariance of the auto-correlation function using the variance of the 29 individual QSO sightlines to evaluate the covariance coefficient, ρ (e.g. Wall & Jenkins 2003). The result is shown in Fig. 4. Strong covariance between adjacent bins is clearly seen with values of $\rho \gtrsim 0.7$. The covariance coefficient generally falls to values of $\rho \lesssim 0.5$ for bins separated by ∼2 or more bins.

We use a simple power-law function of the form $\xi_{\rm F}(s) = (s/s_0)^{-\gamma}$ to fit the clustering measurements. Here, the factor s_0 refers to the (redshift space) correlation length and characterizes the scale at which the probability of finding, in this case, an absorber at a distance s from another absorber is equal to unity (given the power-law profile to ξ). In short, s_0 gives a simple measure of the clustering strength, whereby more highly clustered populations have a longer clustering length.

Fitting our result at scales of $s > 0.9$ Mpc, we find a best-fitting correlation length of $s_0 = 0.081 \pm 0.006\,h^{-1}$ Mpc and a slope of $\gamma = 1.09 \pm 0.04$. The amplitude of the Lyα forest clustering is significantly less than the clustering of the galaxy population at $z \sim 3$ (the corresponding galaxy correlation length is $s_0 \sim 3$–$5\,h^{-1}$ Mpc; e.g. Paper III). Consistent with previous works, the clustering measurement highlights how the Lyα forest is far more uniformly distributed than the galaxy population.

4 DISTRIBUTION OF H I AROUND LBGS

4.1 Lyα–LBG cross-correlation

We now shift our focus to analysing the distribution of Lyα forest absorption line systems around the $z \sim 3$ galaxy population. For this analysis, we measure the cross-correlation function of absorbers around galaxies using our sightline and galaxy survey data. These analyses ultimately allow measurements on the nature of the Lyα forest, offering insights into how it is distributed within large-scale structure and, via redshift-space distortions, the dynamical properties of the forest clouds.

We calculate the Lyα–LBG cross-correlation using the following commonly used form (see e.g. Peebles 1973; Sharp 1979):

$$\xi(s) = \frac{\langle D_{\rm g} D_{\rm T}(s) \rangle}{\langle D_{\rm g} R_{\rm T}(s) \rangle} - 1 \qquad (4)$$

where $\langle D_{\rm g} D_{\rm T}(s) \rangle$ is the weighted galaxy-pixel pair-count at a given separation, s, and $\langle D_{\rm g} R_{\rm T}(s) \rangle$ is the weighted galaxy-random-pixel pair-count as a function of separation. The weighting of a given pair is given by the flux transmission, adjusted for redshift, such that the weighted pair-count is given by $\langle D_{\rm g} D_{\rm T}(s) \rangle = \Sigma \left(T(s)/\overline{T}(z) \right)$.

DLAs, systems with hydrogen column densities $>10^{20}$ cm^{-2}, produce Lyα absorption lines with highly broadened wings. It is therefore important to mask DLA absorption features out of the spectra, since the broad wings would introduce a significant bias when inferring the neutral hydrogen distribution along the sightline from the measured Lyα transmitted flux. We note also that we only use the quasar spectral range between Lyβ and Lyα, since shortward of Lyβ we may not effectively isolate the Lyα forest from the Lyβ forest.

Given the resampling of the high-resolution data, we verify that this has minimal effect on our results by performing the cross-correlation with only the high-resolution QSO spectra, using the non-resampled and re-sampled data sets. The result is shown in Fig. 5. Small deviations are observed in the binned sample at separations of $\lesssim 5\,h^{-1}$ Mpc; however, these changes from the unbinned sample are much smaller than the estimated statistical uncertainties on the points themselves.

The result of the cross-correlation with the 29 QSO spectra is shown in Fig. 6 (filled circles) and follows a power-law form. This measurement of the cross-correlation will inevitably contain some covariance between individual bins, as with all such correlation functions. As with the Lyα forest auto-correlation, we again evaluate the covariance coefficient, ρ, using the individual $\xi(s)$ measurement for each of the 29 background quasars. The result is shown in Fig. 7. We find a covariance coefficient of $\rho \lesssim 0.5$, except for at the high separation bins (i.e. $s \gtrsim 16$), which are strongly correlated with each other ($\rho \approx 0.8$).

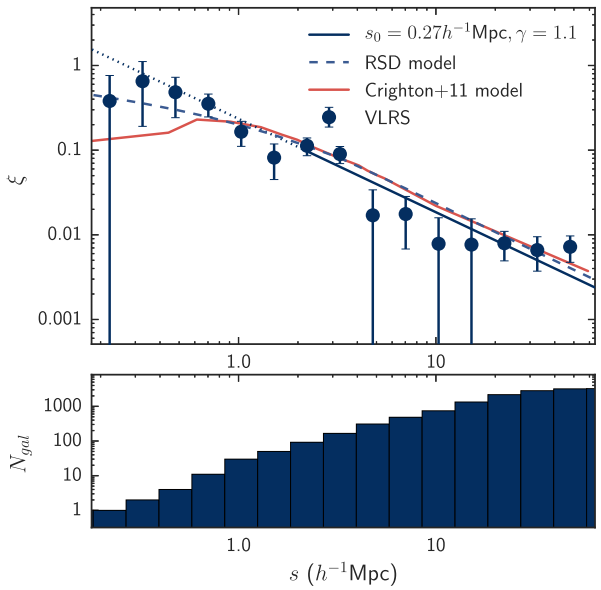

Figure 6. The result of the cross-correlation analysis using the 29 background quasars (filled circles). The solid blue curve shows the best-fitting power-law to the data, whilst the dashed curve shows the $\xi(s)$ model derived from fitting the 2-D cross-correlation in Section 4.3. The red curve shows a model taken from Paper II with a $0.5\,h^{-1}$ Mpc 'transmission spike' in the Lyα transmission profile around star-forming galaxies (and smoothed by a velocity dispersion of 150 km s^{-1}). In the lower panel, we show the numbers of galaxy–absorber pairs as a function of separation, s.

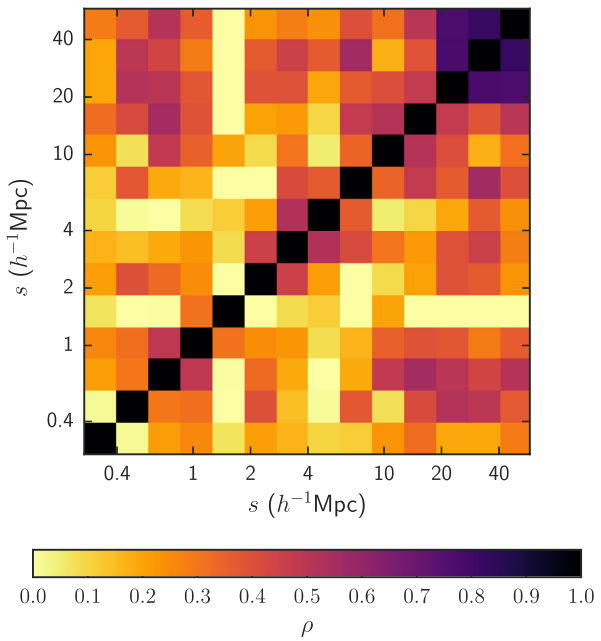

Figure 7. The covariance coefficient measured for the 1D cross-correlation is shown in Fig. 6

Following the same method as we applied for the auto-correlation, we fit a power law to the cross-correlation of the form $\xi(s) = (s/s_0)^{-\gamma}$, where the redshift space correlation length, s_0, is now the distance at which the probability of finding an absorber at a distance, s, from a galaxy is unity. We find a redshift space cross-correlation length of $s_0 = 0.27 \pm 0.14\,h^{-1}$ Mpc and a slope of $\gamma = 1.1 \pm 0.2$ (blue curve in Fig. 6). For reference, we also

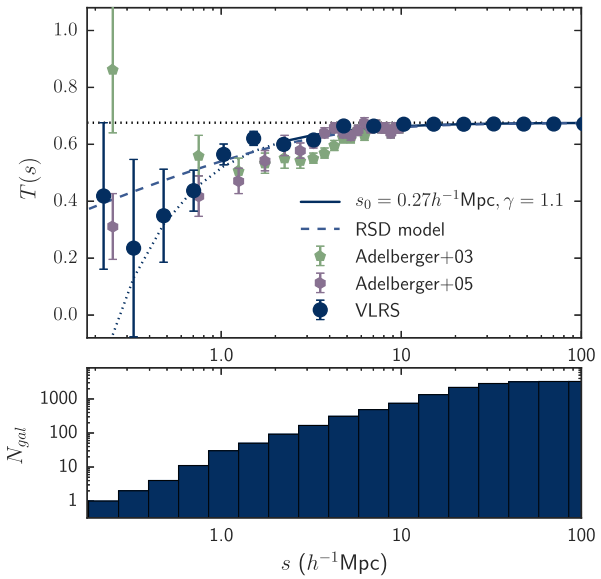

Figure 8. The result of the cross-correlation analysis weighted by transmitted flux based on the 29 background quasars from VLRS with moderate and high-resolution spectra (filled circles). The power law fit to the cross-correlation (also shown in Fig. 6) is again shown by the dashed/solid line, whilst the dashed curve again shows the result from the 2-D cross-correlation fitting presented in Section 4.3. Also shown are the results of Adelberger et al. (2003, filled pentagons) and Adelberger et al. (2005, filled hexagons). The lower panel shows the numbers of galaxy-QSO pairs used in each bin for our analysis.

plot in Fig. 6 the transmission spike model of Paper II (red curve) incorporating a power-law function ($s_0 = 0.3\,h^{-1}$ Mpc and $\gamma = 1.0$) with a transmission spike of width $0.5\,h^{-1}$ Mpc, convolved with a velocity dispersion of 150 km s^{-1}. As with Adelberger et al. (2005), Paper II, Paper IV and Rakic et al. (2012), the data show no evidence for a transmission spike (Adelberger et al. 2003).

For the purposes of comparison with previous works, we present the cross-correlation in terms of the flux transmission, $\langle T(s)\rangle = (1 - \xi(s))\overline{T}(z = 3)$, in Fig. 8. Again our results are shown by the large blue-filled circles. For comparison, the results of Adelberger et al. (2003, filled pentagons) and Adelberger et al. (2005, filled hexagons) are also shown. Consistent with the Adelberger et al. (2005) result at $z \approx 2.4$, our data show strong absorption within $\approx 5\,h^{-1}$ Mpc of the galaxy population. Although reduced absorption is seen in the closest spatial bin, the data point is consistent with monotonically increasing absorption.

There are few previous measurements of the cross-correlation clustering length of galaxies and the Lyα forest, and at redshift $z \approx 3$ we are limited to comparing with our own work in this respect. Paper II showed their data to be consistent with a power law given by $s_0 = 0.3$ and $\gamma = 1.0$, which is consistent within $\approx 1\sigma$ with this latest result. Motivated in part by simulation results, Paper IV attempted to fit a double power law to the cross-correlation, finding $s_0 = 0.49 \pm 0.32$ and $\gamma = 1.47 \pm 0.91$ at scales of $s \gtrsim 1\,h^{-1}$ Mpc, and $s_0 = 0.08 \pm 0.04$ and $\gamma = 0.49 \pm 0.32$ at smaller scales. We find instead that our results are well fit by a single power law, leading to a lower overall clustering length and shallower slope than found using a double power law as in Paper IV. We note, however, that given the large quoted errors on the Paper IV result, our single power law result is in fact consistent with their $s \gtrsim 1\,h^{-1}$ Mpc result. The present more accurate value should be preferred.

Looking to lower redshift, comparisons are made more complex in that measurements are more commonly made using Voigt profile fitting individual lines as opposed to the pixel-based method used here. How these two differently constructed measurements relate to each other has not been fully investigated; however, for completeness we note here the results of Tejos et al. (2014), who performed an analysis of the cross-correlation of galaxies and H I absorption features in QSO sightlines in the redshift range $0 < z < 1$. For their full sample of identified absorbers (with a column density range of $10^{13} < N_{\rm HI} < 10^{17}\,{\rm cm}^{-2}$) they find a correlation length of $r_0 = 1.12 \pm 0.14\,h^{-1}$ Mpc (with a slope of $\gamma = 1.4 \pm 0.1$). On the other hand, for a low-column density subset ($10^{13} < N_{\rm HI} < 10^{14}\,{\rm cm}^{-2}$), they determine a correlation length of $r_0 = 0.14 \pm 0.28\,h^{-1}$ Mpc. Our own measurement (which is based on masking large absorbers from the sightlines) is consistent with the lower column density range of Tejos et al. (2014).

In conclusion, we find consistent correlation lengths with previous studies of the low-density Lyα forest, indicating that the Lyα forest systems are correlated with the star-forming galaxy population, albeit only weakly when compared to the clustering of the galaxies with themselves.

4.2 The effect of absorber strength on the cross-correlation

Prompted by Tejos et al. (2014), we now investigate the presence of any dependence of our results on absorber strength. Whilst the general Lyα forest traces the IGM, as we probe stronger absorbers, current models suggest that we are more likely to be tracing gas structures associated with galactic haloes and filaments. It is interesting therefore to analyse the cross-correlation function as a function of limiting transmitted flux. Indeed, using VLRS data, Pieri et al. (2014) constrained the connection between absorbers and galaxy haloes/the CGM. Taking into account the results presented in their Fig. 4, absorbers with $T_{\rm lim} < 0.33$ are identified as being associated with $\gtrsim$0.2 M$_\star$ LBGs in $\approx$40–50 per cent of occurrences. In addition, they find that such absorbers equate approximately to an average absorber column density of $N_{\rm HI} \sim 10^{15-16.5}\,{\rm cm}^{-2}$ (see their table 1). We note that this equivalence carries with it strong caveats as outlined in Pieri et al. (2014), and use it with caution.

We now use this methodology to investigate the relationship between absorber strength and the galaxy–absorber cross-correlation function. To do so, we first re-sample the QSO spectra used here to match the bin size used in the Pieri et al. (2014) analysis (i.e. to bins of 140 km s^{-1} width). We note that, given this large sampling width, normalized fluxes of $T < 0.33$ in the re-sampled data will consequently not only be tracing strong absorbers but also blended/extended absorbers on the binning scale. We perform the cross-correlation analysis with the re-sampled sightline normalized flux pixels grouped by both limiting normalized flux and discrete bins in normalized flux. The results using the sightline pixels in flux bins of $T \leq 0.33$ ($N_{\rm HI} \sim 10^{15-16.5}\,{\rm cm}^{-2}$), $0.33 \leq T \leq 0.66$ ($N_{\rm HI} \approx 10^{13.5-15}\,{\rm cm}^{-2}$) and $T \geq 0.66$ ($N_{\rm HI} \lesssim 10^{13.5}\,{\rm cm}^{-2}$) are shown in the top panel of Fig. 9, whilst the results using upper limits of $T \leq 0.25$, $T \leq 0.50$ and $T \leq 0.75$ are shown in the lower panel.

A dependence of the clustering amplitude is evident in both sets of results, with lower normalized flux (high column density) corresponding to higher amplitude. This corresponds with the increasing fraction of absorbers predicted to be associated with galaxy haloes as found by Pieri et al. (2014).

We fit the samples with the power-law form already used, taking a fixed power-law slope of $\gamma = 1.1$. The resulting power laws are shown in Fig. 9 as described in the figure legends. We show

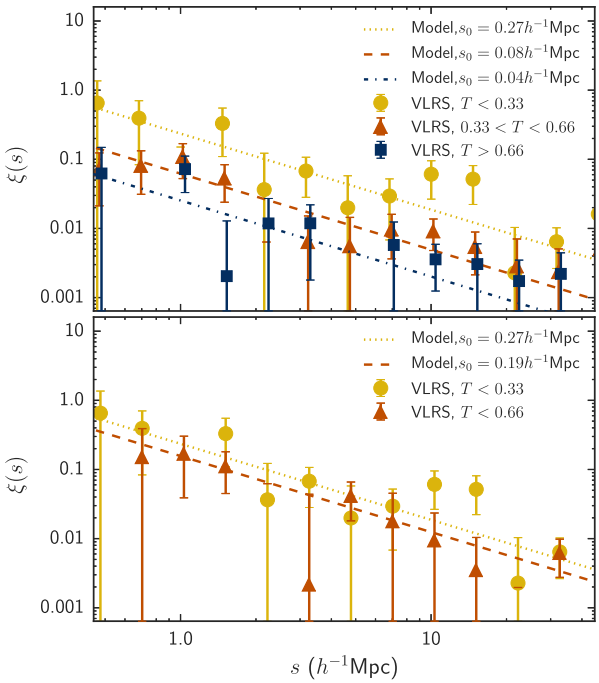

Figure 9. Top panel: Cross-correlation results as a function of transmitted flux bins, with results given for $T > 0.66$ (dark blue squares), $0.33 < T < 0.66$ (orange triangles) and $T < 0.33$ (gold circles). Power-law fits are presented for each with the slope fixed to $\gamma = 1.1$. Lower panel: The equivalent results as a function of a cut in transmitted flux, with results given $T < 0.66$ (orange triangles) and $T < 0.33$ (gold circles). Power-law fits are again presented for each with the slope fixed to $\gamma = 1.1$ in all three cases. The dashed and dotted lines show the fits to the $T < 0.66$ and $T < 0.33$ results, respectively.

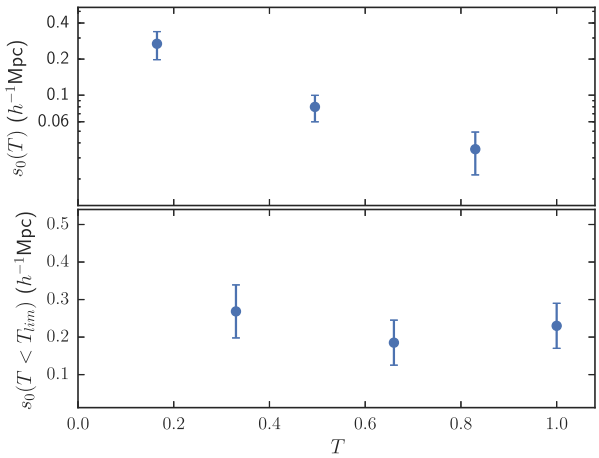

Figure 10. Measured values of the clustering length, s_0, as a function of flux binning (top panel) and a cut in the Lyα transmission flux, $T_{\rm lim}$ (lower pabel). The values correspond to the power-law fits shown in Figs 6 and 9. The data show evidence for a dependence of the clustering length on the cut in Lyα transmission flux, such that stronger absorption systems are more highly correlated with galaxies at $z \sim 3$.

the resulting clustering lengths, s_0, for both the discrete bins (top-panel) and flux limited bins (the bottom panel) in Fig. 10. A trend is observed with the clustering length decreasing with increasing normalized flux in both panels, with the discrete binned sample showing the clearest trend (at the $\approx 4\sigma$ level between the minimum and maximum values). Such a result presents a useful and interesting low-column density corollary to results for high column density

systems such as damped Lyα systems, where clustering is observed to be at a level comparable to the star-forming galaxy population at $z \sim 3$ (e.g. Font-Ribera et al. 2012). It also affirms the connection, employed in Pieri et al. (2014), between absorber strength and mean host halo mass, a parallel to observed relationships between galaxy stellar mass and host-halo mass (e.g. Wake et al. 2011; Bielby et al. 2014; McCracken et al. 2015).

Significant progress has been made with low-redshift studies of the relationship between galaxies and the Lyα forest in recent years with the advent of the Cosmic Origins Spectrograph. Of direct relevance to our high-redshift studies are the results of Tejos et al. (2014), who calculated the projected and two-dimensional cross-correlation functions between Lyα absorbers and galaxies. Given the lower rate of Lyα absorber blending in the forest at lower redshift, Tejos et al. (2014) were able to conduct a complete Voigt profile analysis and perform the cross-correlation using individual absorbers as opposed to using the binned normalized fluxes as we do here. Given their Voigt profile analysis, they are able to cover a much wider range in column density than presented here with our pixel-based method, finding $r_0 = 0.2 \pm 0.4\,h^{-1}$ Mpc for $N < 10^{14}\,\mathrm{cm}^{-2}$ and $r_0 = 3.8 \pm 0.2\,h^{-1}$ Mpc for $N \geq 10^{14}\,\mathrm{cm}^{-2}$ (using their star-forming galaxy sample).

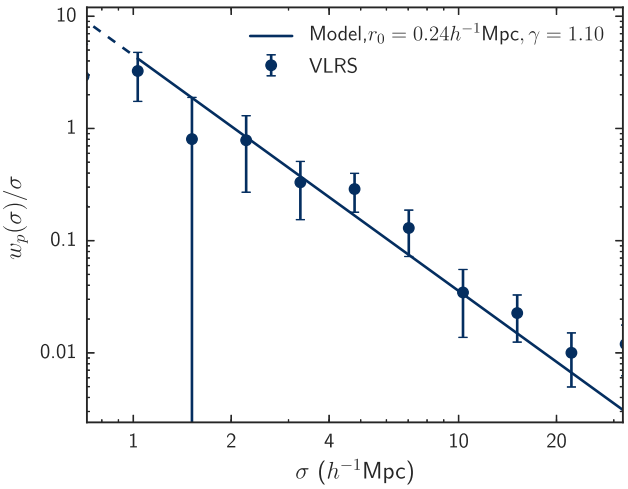

Figure 11. The projected cross-correlation function obtained from the 2-D cross-correlation map (shown in Fig. 12) for the sample of 29 QSO sightlines (solid circle points). The solid line shows a power-law fit to the data (at $1\,h^{-1}$ Mpc $< \sigma < 20\,h^{-1}$ Mpc), given by $r_0 = 0.24 \pm 0.04\,h^{-1}$ Mpc and a fixed slope of $\gamma = 1.1$.

4.3 The projected and 2-D cross-correlation function

In order to isolate and analyse the effects of gas and galaxy dynamics on the cross-correlation, we now turn to the projected and 2-D LBG-Lyα cross-correlation functions. The correlation function, $\xi(s)$, derived in the previous sections contains the imprint of RSDs, such that a 'Kaiser' boost (Kaiser 1987) is observed at separations of more than a few Mpc. This boost increases the amplitude of the clustering measurement. In order to measure the effects of RSD on the correlation function, one requires a model of the underlying 'real-space' clustering, which a model for RSDs we may then apply to, to fit the observed 2-D correlation function and constrain the effects of the RSDs (e.g. Hawkins et al. 2003; Bielby et al. 2013; Tummuangpak et al. 2014).

To derive a model for the real-space custering, we measure the projected correlation function. Simply put, this projects the correlation function along the line of sight, thus removing the imprint of RSDs (which of course apply along the line of sight, but not on transverse, on-sky, separations). In order to distinguish between real and redshift space measurements, we use r to refer to distances in real space (and hence r_0 for the real-space clustering length) and s to refer to distances in redshift space (and hence s_0 for the redshift-space clustering length). Where σ and π are used, these are always real and redshift space separations, respectively (being as they are, the transverse and line-of-sight distances, respectively).

Following Paper IV, we calculate the 2-D cross-correlation function, $\xi(\sigma, \pi)$, and project this along the π direction to derive the projected correlation function, $w_\mathrm{p}(\sigma)$. We calculate $\xi(\sigma, \pi)$ identically to $\xi(s)$ but now on a two-dimensional grid of σ and π. The projection to $w_\mathrm{p}(\sigma)$ is performed by integrating $\xi(\sigma, \pi)$ to some limit in π:

$$w_\mathrm{p} = 2 \int_0^{\pi_\mathrm{max}} \xi(\sigma, \pi) \mathrm{d}\pi \qquad (5)$$

As in the $\xi(s)$ case, we calculate uncertainties on our $\xi(\sigma, \pi)$ and $w_\mathrm{p}(\sigma)$ measurements by performing the analysis on 50 galaxy catalogues each taking the same on-sky spatial distribution as the data, but with randomized redshifts using the measured galaxy redshift distribution as the probability density function.

The resulting projected correlation function is shown in Fig. 11. Again we fit a power law of the form $\xi(r) = (r/r_0)^{-\gamma}$ to the data, excluding separations of $\sigma < 1\,h^{-1}$ Mpc (to avoid small scale effects due to line saturation and line broadening), whilst fitting the data up to a scale of $\sigma < 20\,h^{-1}$ Mpc. This gives a best-fitting clustering length of $r_0 = 0.24 \pm 0.04\,h^{-1}$ Mpc (based on a fixed slope of $\gamma = 1.1$ for consistency with the $\xi(s)$ measurement).

The 2-D cross-correlation result for our sample of galaxies and QSO sightlines is given in Fig. 12. The filled contour map in the left-hand panel shows the 2-D cross-correlation measurement, $\xi(\sigma, \pi)$, the central panel shows the calculated 1σ uncertainties on the measurement, and the right-hand panel shows the number of galaxy–sightline pairs in each bin.

We now proceed to extract the velocity field information from the $\xi(\sigma, \pi)$ measurement. For this we use the velocity field model discussed in Paper IV, which incorporates the effects of small-scale velocity dispersion, characterized by the parameter $\sqrt{\langle w_z^2 \rangle}$, and large-scale infall velocity fields, characterized by the Lyα forest infall parameter β_F (also known as the redshift-space distortion parameter). For a detailed explanation and discussion of the model and these parameters, we refer the reader to Paper IV and Hawkins et al. (2003). In particular, to model the effects of redshift space distortions, we use equation 15 from Paper IV.

For the purposes of fitting the velocity field parameters, we fix the underlying real-space correlation function to that derived from $w_\mathrm{p}(\sigma)$, i.e. a power law with correlation length $r_0 = 0.24\,h^{-1}$ Mpc and a slope of $\gamma = 1.1$. Further, we fix the galaxy infall parameter, β_gal, to the value derived for the LBG sample in Paper III: $\beta_\mathrm{gal} = 0.36$.

The resulting measurements of $\sqrt{\langle w_z^2 \rangle}$ and β_F, derived from a χ^2 minimization with equation 15 from Paper IV, are shown in Fig. 13. We show the confidence limits on the best-fitting values for the infall parameter, $\beta_\mathrm{F} = 1.02 \pm 0.22$, and the velocity dispersion, $\sqrt{\langle w_z^2 \rangle} = 240 \pm 60\,\mathrm{km\,s^{-1}}$. This best-fitting model is shown in the left-hand panel of Fig. 12 by the contour lines, plotted over the measured 2-D cross-correlation function.

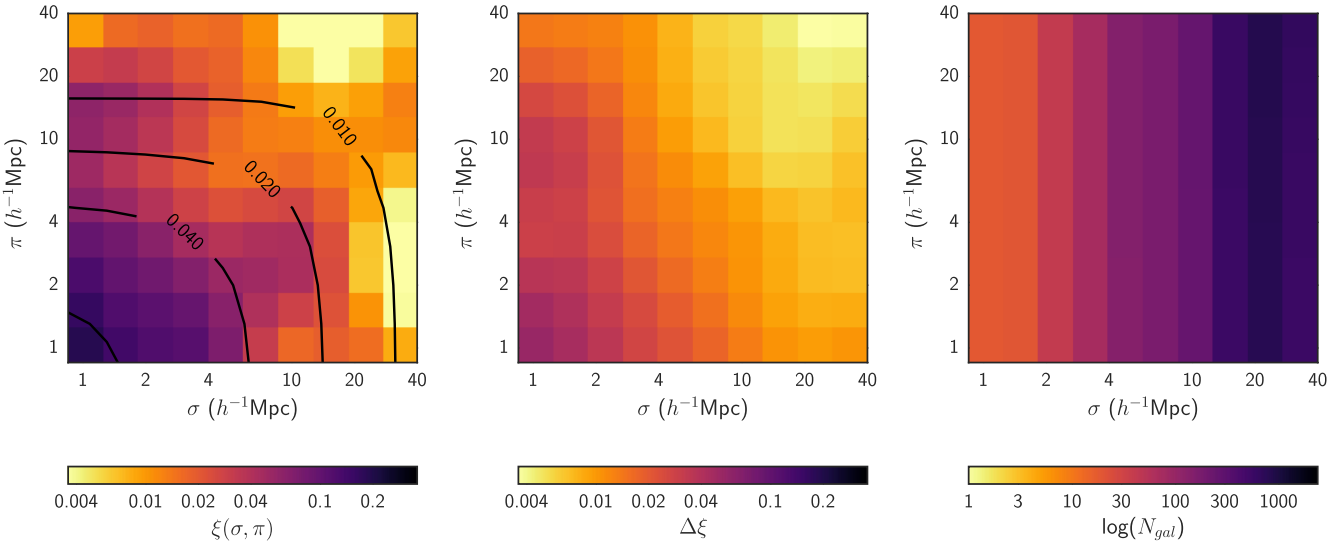

Figure 12. Left-hand side: The $z \approx 3$ 2-D galaxy–galaxy–Lyα cross-correlation function, $\xi(\sigma, \pi)$ for the full (VLT+Keck) sample (shaded contour map). The solid contours show the best-fitting RSD model with $\beta_F = 1.02 \pm 0.22$ and $\sqrt{\langle w_z^2 \rangle} = 240 \pm 60$. Centre: The estimated uncertainties on $\xi(\sigma, \pi)$, calculated based on 50 random realisations of the galaxy catalogue. Right-hand side: The number of galaxies used in each bin.

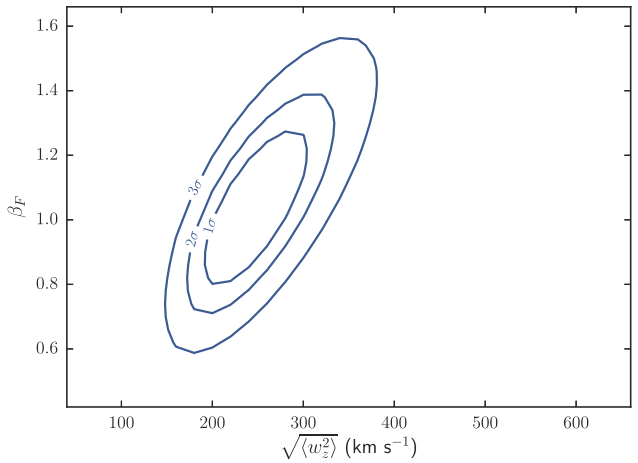

Figure 13. Confidence contours on fitting the 2-D cross-correlation function for the Lyα forest infall parameter, β_F and the velocity dispersion, $\sqrt{\langle w_z^2 \rangle}$, for the combined VLT+Keck data sample. The fit assumes an underlying power law form to the intrinsic cross-correlation function with correlation length $r_0 = 0.24\,h^{-1}$ Mpc and a slope of $\gamma = 1.1$.

5 DISCUSSION

5.1 Absorber dynamics

5.1.1 Infall/RSD parameter

Few statistical or quantitative measures of the Lyα forest absorber–galaxy dynamics have been made at any redshift, whilst model predictions are also relatively sparse. The primary route to constraining the dynamics of the Lyα forest has instead been via the Lyα forest auto-correlation function or power spectrum (e.g. McDonald 2003; Slosar et al. 2011; Seljak 2012). Rakic et al. (2013) showed evidence of large-scale infall in their cross-correlation analysis, but were limited to scales of $\lesssim 7\,h^{-1}$ Mpc and did not present any constraints on the large-scale dynamics themselves.

Paper IV presented an analysis of these dynamics with both a subset of the data used here and a subvolume snapshot of the GIMIC simulation. Using equivalent analyses to those presented here, the data gave constraints on the infall parameter of $\beta_F = 0.33^{+0.23}_{-0.33}$, whilst the analysis of the simulation predicted a somewhat higher value of $\beta_F = 0.51 \pm 0.12$ (although the authors note that this simulated measurement may be affected by the relatively small size of the simulated volume). We have now improved significantly on the previous result by adding the X-Shooter data to the analysis, in particular improving our constraints on the underlying input (real-space) model given by the $w_p(\sigma)$ measurement. Given the improved constraints on r_0 from $w_p(\sigma)$ and the enhanced signal on the 2-D cross-correlation signal, we find that our measurement of the infall parameter is higher than that suggested in Paper IV.

Measurement of the large-scale dynamics of the Lyα forest from the auto-correlation function has largely been based on BOSS, with original results from BOSS giving constraints of $\beta = 0.8 \pm 0.2$ (Slosar et al. 2011). Using the Lyα forest measurements of BOSS Data Release 11 ($z = 2.3$), Delubac et al. (2015) and Blomqvist et al. (2015) find best-fitting measurements of $\beta_F = 1.50 \pm 0.47$ and $\beta_F = 1.39^{+0.11}_{-0.10}$, respectively. Closer in method to our own work, Font-Ribera et al. (2013) derive β_F from the cross-correlation between the Lyα forest and the BOSS QSO sample at $z \approx 2.3$, finding $\beta_F = 1.10^{+0.17}_{-0.15}$. Most models predict some evolution in the infall parameter between the BOSS redshift and that of our analysis. Indeed, the simulation results of Arinyo-i-Prats et al. (2015) would suggest a factor of $\beta_F(z = 2.8)/\beta_F(z = 2.3) \approx 0.92$, translating the most recent BOSS results to $\beta_F \approx 1.28$–1.38 at $z = 2.8$. Our own result lies somewhere between the early and more recent BOSS results and, at $\beta_F = 1.02 \pm 0.22$, favours a marginally ($\approx 1\sigma$) weaker level of IGM gas infall compared to the most recent results.

Paper IV also compared their result with the simulations of McDonald (2003). Based on several re-simulations varying the simulation parameters, McDonald (2003) predicts a Lyα infall parameter of $\beta_F = 1.58 \pm 0.05$ (or $\beta_F = 1.45$ taking into account redshift evolution). Although the tension between this and our own results is now smaller, we still find a significantly lower infall parameter than predicted by McDonald (2003) at the $\approx 3\sigma$ level.

Interestingly, more recent simulations place the predictions of McDonald (2003) towards the higher end of published predictions. Seljak (2012) noted that the RSD parameter could feasibly lie within the range $0.5 \lesssim \beta_{\rm F} \lesssim 1.5$, assuming a realistic range of bias parameters for the Lyα forest. Finally, Lochhaas et al. (2016) use the Lyα Mass Association Scheme (LyMAS) and predict $\beta_{\rm F} = 0.970 \pm 0.016$ at $z = 2.5$, highlighting the range of predictions made via simulations. Once redshift evolution is taken into account, this lies $\approx 1\sigma$ below our observations.

5.1.2 *Small-scale velocity dispersion*

The total velocity dispersion imprinted upon the 2-D cross-correlation function of $\sqrt{\langle w_z^2 \rangle} = 240 \pm 60\,{\rm km\,s^{-1}}$ comprises three primary components: the measurement uncertainty on the galaxy redshifts; the thermal broadening of the sightline absorption lines; and the intrinsic velocity dispersion between galaxies and Lyα clouds. For the VLT VIMOS galaxy data, the galaxy redshift errors are $\approx 350\,{\rm km\,s^{-1}}$, whilst the Keck LRIS data used for nine of the sightlines provide marginally more accurate redshifts at $\approx 250\,{\rm km\,s^{-1}}$. From this, it is evident that our result for the velocity dispersion is dominated by the galaxy redshift errors. Indeed the thermal broadening on the absorption lines is $\approx 70\,{\rm km\,s^{-1}}$ (Paper II), which, if we combine in quadrature with the average of the instrumental errors on the galaxy redshifts, gives $\sqrt{\langle w_z^2 \rangle} = 310\,{\rm km\,s^{-1}}$, i.e. $\approx 1\sigma$ larger than the best-fitting measurement of $\sqrt{\langle w_z^2 \rangle}$ from the $\xi(\sigma, \pi)$ measurement. Combining this estimate of the instrumental plus thermal broadening effects with our results in Fig. 13, we calculate a 3σ upper limit on the intrinsic LBG-Lyα velocity dispersion of $\sqrt{\langle w_z^2 \rangle} < 220$ km s^{-1}.

Ultimately, we require observations with reduced velocity uncertainties to more closely analyse the intrinsic velocity dispersion, however this approximate upper limit still offers some insights and opportunity for comparison with other works at both low and high redshifts. Ryan-Weber (2006) measured a large finger-of-god velocity dispersion effect in the cross-correlation between absorbers and galaxies at $z \sim 0$, over scales of $\approx$400–600 km s^{-1}. Similarly, Rakic et al. (2012) and Turner et al. (2014) show, using the KBSS data at $z \sim 2.3$, elongations extending to $\approx 200\,{\rm km\,s^{-1}}$ along the line of sight. Conversely, Tejos et al. (2014) claim an upper limit of $\lesssim 120\,{\rm km\,s^{-1}}$ when analysing the cross-correlation of galaxies and absorbers at $0 < z < 1$.

5.2 Absorbers tracing the underlying dark matter distribution

We now consider the relationship between the Lyα forest, the galaxy population and the underlying dark matter distribution. Following the example of Adelberger et al. (2003) and Tejos et al. (2014), we use the Cauchy–Schwarz inequality to evaluate the connection between the LBG population and the Lyα forest. The Cauchy–Schwarz inequality takes the form:

$$\xi_{\rm ag}^2 \leq \xi_{\rm aa}\xi_{\rm gg} \tag{6}$$

where $\xi_{\rm ag}$ is the cross-correlation between absorbers and galaxies, $\xi_{\rm aa}$ is the auto-correlation between absorbers and $\xi_{\rm gg}$ is the galaxy–galaxy auto-correlation. If the two sides of this equation are equal, then it follows that the two populations being evaluated trace the same dark matter structure and the difference in the clustering biases can be used to surmise the relative masses of dark-matter haloes that the populations trace within the overall matter structure. On the other hand, if the equality does not hold, i.e. $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) < 1$, then this can potentially provide insights into the baryonic physics affecting the two populations (assuming that the standard cosmological paradigm is correct).

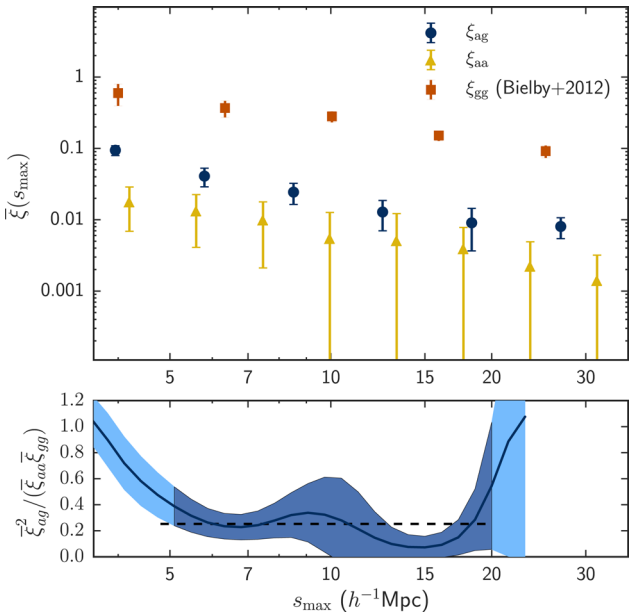

Figure 14. Top panel: Integrated clustering functions for the LBG-Lyα forest cross-correlation; the LBG auto-correlation (taken from Paper III); and the Lyα forest auto-correlation. Lower panel: The Cauchy–Schwarz ratio calculated from the three correlation functions. The dark shaded region shows the 1σ range in the ratio over scales of $5 \leq s_{\rm max} \leq 20$, representing the linear regime. The dashed line shows the median value in this range.

We evaluate the Cauchy–Schwarz inequality by calculating the integrated clustering functions, $\overline{\xi}(s_{\rm max})$, of each correlation function, which is given by:

$$\overline{\xi}(s_{\rm max}) = \int_0^{s_{\rm max}} \xi(s) s^2 {\rm d}s \tag{7}$$

The integrated clustering functions and the Cauchy–Schwarz ratio are shown in Fig. 14. Taking into account the range $5 \leq s_{\rm max} \leq 20\,h^{-1}$ Mpc (i.e. large enough scales to be in the linear regime, whilst small enough that the uncertainties in the measurements are still relatively low), we find a median value of $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) = 0.25 \pm 0.14$. The minimum value in this range is using $s_{\rm max} = 15\,h^{-1}$ Mpc, which gives $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) = 0.07 \pm 0.09$, whilst the maximum is at $s_{\rm max} = 20\,h^{-1}$ Mpc, which gives $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) = 0.55 \pm 0.49$. There is significant variation even in this range then, however the $\xi_{\rm ag}^2 < \xi_{\rm aa}\xi_{\rm gg}$ remains less than unity, maintaining the relationship as an inequality.

That the ratio is less than unity is a very strong indication that the underlying baryonic matter distributions giving rise to the Lyα forest absorption systems and galaxies are not linearly dependent. This adds to results at low redshift where weak systems such as populate the forest give $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) < 1$, whilst stronger absorbers ($N_{\rm HI} \gtrsim 10^{14}\,{\rm cm^{-2}}$) are consistent with $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) = 1$ (Tejos et al. 2014).

6 CONCLUSIONS

We have used the VLRS galaxy data set in conjunction with high-quality moderate- and high-resolution background quasar spectra to investigate the relationship between galaxies and the H I gas of

the IGM. The analysis incorporates 29 $z \gtrsim 3$ quasars, 15 of which have been observed with the VLT X-Shooter instrument.

We have calculated the Lyα auto-correlation function using our full sample using both line-of-sight and cross line-of-sight ($\sigma \gtrsim 6\,h^{-1}$ Mpc) data and fit a power-law form, finding a clustering length of $s_0 = 0.081 \pm 0.006\,h^{-1}$ Mpc and slope of $\gamma = 1.09 \pm 0.04$.

Using a large spectroscopic sample of LBGs at $z \approx 3$, we determine the LBG–Lyα cross-correlation function, $\xi(s)$. As with the Lyα auto-correlation, we fit the data with a power law, finding a clustering length of $s_0 = 0.27 \pm 0.14\,h^{-1}$ Mpc and slope of $\gamma = 1.1 \pm 0.2$, improving on the accuracy of our analyses presented in Paper II and Paper IV. These auto- and cross-correlation results highlight the weak clustering of the Lyα forest both in itself and in relation to the galaxy population (when compared to metal absorption line systems/optically thick systems and galaxies themselves). Further to this, we calculate the LBG–Lyα cross-correlation function as a function of normalized flux, T (a proxy for column density), finding a significant anti-correlation of the resulting clustering lengths with T. This shows that higher density H I absorbers are more strongly clustered around galaxies at $z \sim 3$, whilst low-density absorbers are only weakly clustered with the galaxy population – i.e. stronger absorbers trace higher mass haloes on average.

Further to the one-dimensional cross-correlation analysis, we calculate the projected and two-dimensional LBG–Lyα cross-correlation functions. Using the projected correlation function, we constrain the real-space correlation length to be $r_0 = 0.24 \pm 0.04\,h^{-1}$ Mpc (assuming a fixed slope of $\gamma = 1.1$ based on the $\xi(s)$ result). Combining this result with our measurement of $\xi(\sigma, \pi)$, we constrain the dynamical properties of the LBG-H I density field as probed by the galaxy survey-QSO sightline data. We find $\beta_{\rm F} = 1.02 \pm 0.22$ and $\sqrt{\langle w_z^2 \rangle} = 240 \pm 60\,{\rm km\,s^{-1}}$. This presents a new and clear detection of the large–scale infall of gas towards high–density regions within a large–scale structure. Our measurement of the velocity dispersion between the galaxy and gas components is consistent with the uncertainties on the galaxy redshift measurements, but does give a 3σ upper limit on the intrinsic LBG-Lyα velocity dispersion of $\sqrt{\langle w_z^2 \rangle} < 220$ km s^{-1}, similar to our previous result in Paper IV.

The combination of our auto-correlation and cross-correlation results, along with our previous results for the LBG auto-correlation function, allows us to evaluate the Cauchy–Schwartz inequality, which we find to be significantly below unity: $\xi_{\rm ag}^2/(\xi_{\rm aa}\xi_{\rm gg}) = 0.25 \pm 0.14$. Combined with the weak cross-clustering signal, this highlights how Lyα forest absorbers do not linearly follow the density profile traced by the galaxy population.

ACKNOWLEDGEMENTS

Based on observations collected at the European Organization for Astronomical Research in the Southern hemisphere under ESO programmes 075.A-0683, 077.A-0612, 079.A-0442, 081.A-0474, 082.A-0494, 085.A-0327 and 087.A-0906. Also based on data obtained with the NOAO Mayall 4 m Telescope at Kitt Peak National Observatory, USA (programme ID: 06A-0133) and the NOAO Blanco 4 m Telescope at Cerro Tololo Inter-American Observatory, Chile (programme IDs: 03B-0162, 04B-0022).

RMB, TS and SLM are funded by the STFC research grant ST/L00075X/1. NHMC thanks the Australian Research Council for *Discovery Project* grant DP130100568, which in part supported this work. DM is supported by the BASAL CATA grant PFB-06 and the Millennium Institute for Astrophysics MAS grant IC120009. LI acknowledges Conicyt grants Basal-CATA PFB-06/2007 and Anillo ACT1417.

This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

REFERENCES

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45
Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005, ApJ, 629, 636
Arinyo-i-Prats A., Miralda-Escudé J., Viel M., Cen R., 2015, J. Cosmol. Astropart. Phys., 12, 017
Bahcall J. N., Jannuzi B. T., Schneider D. P., Hartig G. F., Green R. F., 1992, ApJ, 397, 68
Bi H., Davidsen A. F., 1997, ApJ, 479, 523
Bielby R. M. et al., 2011, MNRAS, 414, 2 (Paper I)
Bielby R. et al., 2013, MNRAS, 430, 425 (Paper III)
Bielby R. M. et al., 2014, A&A, 568, A24
Bielby R. M. et al., 2016, MNRAS, 456, 4061 (Paper V)
Bielby R., Crighton N. H. M., Fumagalli M., Morris S. L., Stott J. P., Tejos N., Cantalupo S., 2017, MNRAS, 468, 1373
Blomqvist M. et al., 2015, J. Cosmol. Astropart. Phys., 11, 034
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Carswell R. F., Whelan J. A. J., Smith M. G., Boksenberg A., Tytler D., 1982, MNRAS, 198, 91
Chen H.-W., Prochaska J. X., 2000, ApJ, 543, L9
Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 1998, ApJ, 498, 77
Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 2001, ApJ, 559, 654
Crighton N. H. M., Morris S. L., Bechtold J., Crain R. A., Jannuzi B. T., Shone A., Theuns T., 2010, MNRAS, 402, 1273
Crighton N. H. M. et al., 2011, MNRAS, 414, 28 (Paper II)
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 11
Delubac T. et al., 2015, A&A, 574, A59
Font-Ribera A. et al., 2012, J. Cosmol. Astropart. Phys., 11, 059
Font-Ribera A. et al., 2013, J. Cosmol. Astropart. Phys., 5, 018
Hawkins E. et al., 2003, MNRAS, 346, 78
Heckman T. M., Armus L., Miley G. K., 1990, ApJS, 74, 833
Hoffman G. L., Lu N. Y., Salpeter E. E., Connell B. M., 1999, AJ, 117, 811
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Hopkins P. F., Quataert E., Murray N., 2012, MNRAS, 421, 3522
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Kaiser N., 1987, MNRAS, 227, 1
Lanzetta K. M., Bowen D. V., Tytler D., Webb J. K., 1995, ApJ, 442, 538
Lehnert M. D., Heckman T. M., 1996, ApJ, 472, 546
Lochhaas C. et al., 2016, MNRAS, 461, 4353
McCracken H. J. et al., 2015, MNRAS, 449, 901
McDonald P., 2003, ApJ, 585, 34
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P., Miralda-Escudé J., Cen R., 2002, ApJ, 580, 42
Morris S. L., Jannuzi B. T., 2006, MNRAS, 367, 1261
Morris S. L., van den Bergh S., 1994, ApJ, 427, 696
Morris S. L., Weymann R. J., Dressler A., McCarthy P. J., Smith B. A., Terrile R. J., Giovanelli R., Irwin M., 1993, ApJ, 419, 524
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325
Pachat S., Narayanan A., Muzahid S., Khaire V., Srianand R., Wakker B. P., Savage B. D., 2016, MNRAS, 458, 733
Peebles P. J. E., 1973, ApJ, 185, 413

Pieri M. M. et al., 2014, MNRAS, 441, 1718
Rakic O., Schaye J., Steidel C. C., Rudie G. C., 2012, ApJ, 751, 94
Rakic O., Schaye J., Steidel C. C., Booth C. M., Dalla Vecchia C., Rudie G. C., 2013, MNRAS, 433, 3103
Rudie G. C. et al., 2012, ApJ, 750, 67
Ryan-Weber E. V., 2006, MNRAS, 367, 1251
Schaye J. et al., 2010, MNRAS, 402, 1536
Seljak U., 2012, J. Cosmol. Astropart. Phys., 3, 004
Sharp N. A., 1979, A&A, 74, 308
Shen S., Wadsley J., Stinson G., 2010, MNRAS, 407, 1581
Shone A. M., Morris S. L., Crighton N., Wilman R. J., 2010, MNRAS, 402, 2520
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Slosar A. et al., 2011, J. Cosmol. Astropart. Phys., 9, 001
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728
Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, ApJ, 736, 160
Tejos N., Morris S. L., Crighton N. H. M., Theuns T., Altay G., Finn C. W., 2012, MNRAS, 425, 245
Tejos N. et al., 2014, MNRAS, 437, 2017
Tejos N. et al., 2016, MNRAS, 455, 2662
Tummuangpak P., Bielby R. M., Shanks T., Theuns T., Crighton N. H. M., Francke H., Infante L., 2014, MNRAS, 442, 2094 (Paper IV)
Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, MNRAS, 445, 794
Wake D. A. et al., 2011, ApJ, 728, 46
Wall J., Jenkins C., 2003, Practical Statistics for Astronomers. Cambridge Observing Handbooks for Research Astronomers, Cambridge Univ. Press, Cambridge
Wilman R. J., Gerssen J., Bower R. G., Morris S. L., Bacon R., de Zeeuw P. T., Davies R. L., 2005, Nature, 436, 227
Wilman R. J., Morris S. L., Jannuzi B. T., Davé R., Shone A. M., 2007, MNRAS, 375, 735
Wisotzki L. et al., 2016, A&A, 587, A98
Young P. J., Sargent W. L. W., Boksenberg A., Carswell R. F., Whelan J. A. J., 1979, ApJ, 229, 891

This paper has been typeset from a TEX/LATEX file prepared by the author.